IN THE CLAIMS

1. (Currently Amended) A method of processing an information signal containing content presented in accordance with at least one modality, the method comprising the steps of:

obtaining the information signal, wherein the information signal represents speech obtained in accordance with one of a voice message system and a real-time phone conversation;

performing content detection on the information signal to detect whether the information signal includes particular content presented in accordance with the at least one modality; and

automatically generating a control signal; as a direct result of detection of when the particular content is detected without a need for a selective input by a user contemporaneous with the detection of the content, for use in automatically controlling at least one of a rendering property of the particular content and automatic implementation of a specific action relating to the particular detected content.

2. (Original) The method of claim 1, wherein the at least one modality in which the content in the information signal is presented is audio-based.

3. (Original) The method of claim 1, wherein the at least one modality in which the content in the information signal is presented is video-based.

4. (Original) The method of claim 1, wherein the at least one modality in which the content in the information signal is presented is audio-based and video-based.

5. (Original) The method of claim 1, wherein the controlled rendering property is a presentation speed of the particular content.

6. (Original) The method of claim 5, wherein the presentation speed is controlled in accordance with detection of specific content classes in the information signal.

7. (Original) The method of claim 6, wherein a specific content class comprises one of numbers, names and addresses.

8. (Original) The method of claim 5, wherein the presentation speed of the particular content is at least one of slowed down and sped up.

9. (Original) The method of claim 5, wherein the presentation speed of the particular content is slowed down from an initial sped-up presentation speed.

10. (Original) The method of claim 1, further comprising the step of providing a user interface for a user to control at least one of the rendering property of the particular content and the implementation of the specific action relating to the particular content.

11. (Original) The method of claim 1, further comprising the step of marking at least a portion of the information signal in response to a user input such that the content detection step is performed on the marked portion of the information signal to detect whether the marked portion of the information signal includes the particular content.

12. (Original) The method of claim 1, further comprising the step of storing the particular content when detected in the information signal.

13. through 27. (Cancel).

28. (Currently Amended) A method of processing an information signal containing content presented in accordance with at least one modality, the method comprising the steps of:

obtaining the information signal, wherein the information signal represents speech obtained in accordance with one of a voice message system and a real-time phone conversation;

marking at least a portion of the information signal in response to a user input;

performing content detection on the at least a portion of the information signal to detect whether the marked portion of the information signal includes desired content presented in accordance with the at least one modality; and

at least one of storing and utilizing the desired content in a subsequent application when detected in the information signal.

29. (Cancel).

30. (Currently Amended) A method of processing an information signal containing content presented in accordance with at least two modalities, the method comprising the steps of:

obtaining the information signal;

separating the information signal into a first signal including information in one of the two modalities and a second signal including information in the other of the two modalities;

performing content detection on the first signal to detect whether the first signal includes particular content presented in accordance with the one modality;

performing content detection on the second signal to detect whether the second signal includes particular content presented in accordance with the other modality;

combining results associated with the content detection steps; and

automatically generating a control signal, as a direct result of detection of when at least a portion of the particular content is detected in accordance with at least one of the content detection steps without a need for a selective input by a user contemporaneous with the detection of the content, for use in automatically controlling at least one of a rendering property of the particular content and automatic implementation of a specific action relating to the particular detected content.

31. (Original) The method of claim 30, wherein the two modalities are video and audio.

32. (Original) The method of claim 31, wherein the content detection step performed on the video signal is optical character recognition and the content detection step performed on the audio signal is speech recognition.

33. through 44. (Cancel).

45. (Currently Amended) Apparatus for processing an information signal containing content presented in accordance with at least one modality, the apparatus comprising:

at least one processor operative to: (i) obtain the information signal, wherein the information signal represents speech obtained in accordance with one of a voice message system and a real-time phone conversation; (ii) perform content detection on the information signal to detect whether the information signal includes particular content presented in accordance with the at least one modality; and (iii) automatically generate a control signal, as a direct result of detection of when the particular content is detected without a need for a selective input by a user contemporaneous with the detection of the content, for use in automatically controlling at least one of a rendering property of the particular content and automatic implementation of a specific action relating to the particular detected content.

46. (Cancel).

47. (Cancel).

48. (Currently Amended) Apparatus for processing an information signal containing content presented in accordance with at least one modality, the apparatus comprising:

at least one processor operative to: (i) obtain the information signal, wherein the information signal represents speech obtained in accordance with one of a voice message system and a real-time phone conversation; (ii) mark at least a portion of the information signal in response to a user input; (iii) perform content detection on the at least a portion of the information signal to detect whether

the marked portion of the information signal includes desired content presented in accordance with the at least one modality; and (iv) at least one of store and utilize the desired content in a subsequent application when detected in the information signal.

49. (Cancel).

50. (Currently Amended) Apparatus for processing an information signal containing content presented in accordance with at least two modalities, the apparatus comprising:

at least one processor operative to: (i) obtain the information signal; (ii) separate the information signal into a first signal including information in one of the two modalities and a second signal including information in the other of the two modalities; (iii) perform content detection on the first signal to detect whether the first signal includes particular content presented in accordance with the one modality; (iv) perform content detection on the second signal to detect whether the second signal includes particular content presented in accordance with the other modality; (v) combine results associated with the content detection steps; and (vi) <u>automatically</u> generate a control signal, as a direct result of detection of <s>when</s> at least a portion of the particular content <s>is detected</s> in accordance with at least one of the content detection steps <u>without a need for a selective input by a user</u>, for use in <u>automatically</u> controlling at least one of a rendering property of the particular content and <u>automatic</u> implementation of a specific action relating to the particular <u>detected</u> content.

51. through 53. (Cancel).